| *Motion Imagery Standards Board*<br>**Engineering Guideline**<br><br>**Delivery of Low Bandwidth Motion Imagery** | **MISB EG 0803**<br><br>**24 April 2008** |
|---|---|

## 1 Scope

This Motion Imagery Standards Board (MISB) Engineering Guideline (EG) provides information for creating and distributing Motion Imagery (MI) over low bandwidth channels.

## 2 References


[1] MISP 4.5, *Motion Imagery Standards Profile*

[2] MISB EG 0104, *Predator UAV Basic Universal Metadata Set, December 2006*

[3] MISB EG 0601, *UAS Datalink Local Metadata Set, May 2008*

[4] MISB RP 0701, *Common Metadata System: Structure, August 2007*

[5] ISO/IEC 13818-1, *Information Technology – Generic coding of moving pictures and associated audio information: Systems, 2000*

[6] ITU-T Rec. H.222, Amendment 3, 2004: *Transport of AVC data over ISO/IEC 138181/H.222.0 for MPEG2 TS containment for MPEG4 AVC*

[7] RFC 3550 *RTP: A Transport Protocol for Real-Time Applications, July 2003*

[8] RFC 3984 *RTP Payload Format for H.264 Video, February 2005*

[9] RFC 2326, *Real Time Streaming Protocol (RTSP)*, April 1998

[10] RFC 2327, *SDP: Session Description Protocol*, April 1998

[11] RFC 2616, *Hypertext Transfer Protocol -- HTTP/1.1, 1999*

[12] ISO/IEC 13818-1:2000/AMD 1: 2003, Information technology -- *Generic coding of moving pictures and associated audio information: Systems, AMENDMENT 1: carriage of metadata over ISO/IEC 13818-1 streams.*


## 3 Acronyms

| | |
|---|---|
| AAC | Advanced Audio Coding |
| FTP | File Transfer Protocol |
| HTTP | Hypertext Transfer Protocol |

| | |
|---|---|
| IP | Internet Protocol |
| RTP | Real Time Protocol |
| RTCP | Real Time Control Protocol |
| RTSP | Real Time Streaming Protocol |
| SAP | Session Announcement Protocol |
| SBR | Spectral Band Replication |
| SDP | Session Description Protocol |
| TCP | Transmission Control Protocol |
| Xon2 | Compression Type “X” carried over MPEP2 Transport Stream |
| UDP | User Datagram Protocol |
| URL | Uniform Resource Locator |

# 4 Introduction

The motion imagery produced at the Command and Control (C2) station of a motion imagery collection system (collection platform and control station) is considered the exploitation quality motion imagery for the system. The typical exploitation quality motion imagery is generated at MISP MISM level 3 [1] and greater. The motion imagery produced at these levels can exceed the bandwidth of the network delivering motion imagery to users at the edges. To meet network constraints motion imagery asset tradeoffs are required; either the imagery must be reduced in spatial resolution, decreased in temporal rate, or reduced in fidelity. In some cases a combination of reductions is necessary. Alternative measures to meet bandwidth constraints include reducing the quantity of metadata, and transcoding the imagery using different compression. The video quality of a reduced-bandwidth motion imagery asset is typically in the range of MISP MISM levels 0 to 2 [1]. Because salient features in the original imagery may be lost in such highly compressed imagery, it should not be relied on for exploitation and targeting purposes.

Characterizing a channel as “low bandwidth” usually suggests a narrow channel or pipe, where information flow will not support real-time motion imagery, such as video. But low bandwidth also can include higher-capacity channels where interference may be high causing a reduced data throughput, or excessive network traffic incurs collisions and lost data. What is relevant is the overall data throughput rather than a high versus low bandwidth designation. Knowledge of network conditions is thus important in determining what technologies are best choices to achieve a desired level of performance.

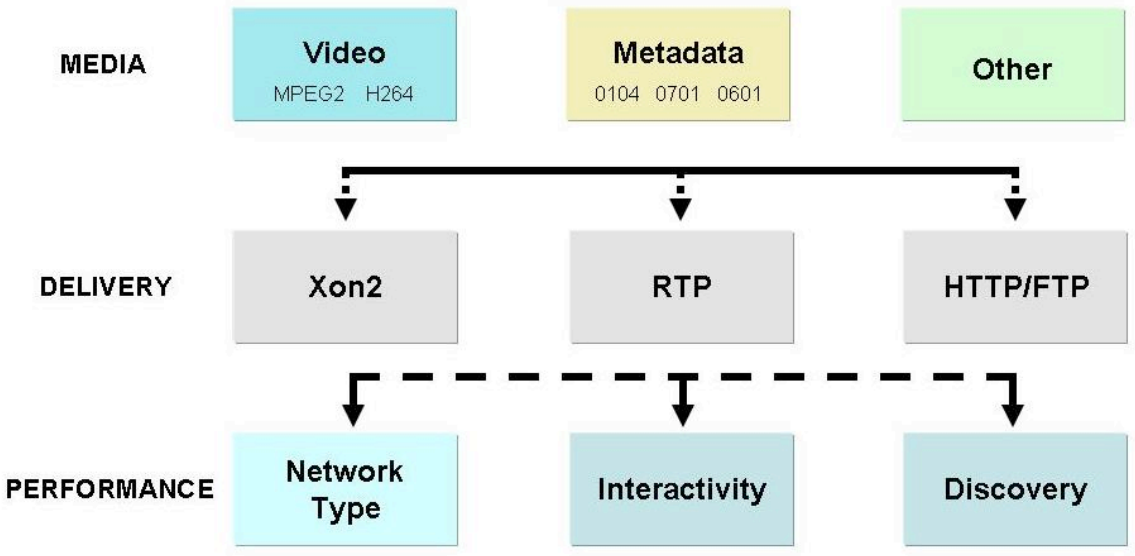
This Engineering Guideline is an aide for choosing technologies that meet design criteria for a given level of user performance within the MISB portfolio of adopted standards and practices.

# 5 Low Bandwidth Motion Imagery Components

Design objectives and criteria will suggest some particular combination of media types, delivery protocol, and distribution performance that will shape the user experience (see

figure below). The network type (i.e. IP, RF link, etc.), the level of interactivity with the content (i.e. TiVo functions, real-time, etc.), and a means for discovery of content (pull versus push) are all factors in choosing an optimal delivery protocol for an application. Xon2, RTP, and HTTP/FTP are delivery protocols for the video, metadata, and other media.

MISB-compliant video is either MPEG2 or H.264 compressed and associated with KLV metadata structured according to Engineering Guidelines EG 0104 [2], EG 0601 [3], and RP 0701 [4]. Audio and other annotation data may be included, but is currently not specified here.




**Elements That Shape a User Experience**

## *5.1 Video Encoding*

The approved video encoding formats for low-bandwidth motion imagery is specified in the MISP RECOMMENDED PRACTICE 9720e–MISM, Low Bandwidth Motion Imagery Levels L1.2-L1.0. The MISB approved compression for these MISM levels is H.264/AVC, which supports data rates between 384 and 56 Kbits/sec.

## *5.2 Video Distribution Protocols*

The approved distribution protocols for real-time delivery of low-bandwidth motion imagery are Xon2 and RTP (Real Time Protocol). The choice in delivery protocol is largely a matter of network type, desired interactivity, performance, and to a lesser degree the discovery means for available content.

### 5.2.1 Xon2

The Xon2 protocol is the MPEG-2 Transport Stream (TS), but indicated as “X” on “2” because it can support other compression technologies, such as H.264/AVC on MPEG2 TS. Together, the MPEG2 Transport Stream protocol [5, 6] and the MISB standards and recommendations define the rules for transporting video, audio, metadata, and other supporting data within the stream. MPEG2 TS carries all media components (video, metadata, other) associated with a collect within one unified transport package.

#### 5.2.1.1 Network Type

MPEG2 TS is designed for constant delay networks and robustness in the presence of channel noise and interference, such as is typical in broadcast application. Thus, for transmission mediums that are point-to-point circuit-based, such as an RF link, MPEG2 TS is a preferred protocol. MPEG2 TS is frequently used to carry multimedia over *packet-based* networks, where best performance is observed when these networks are dedicated with little conflicting traffic and few network hops. The DoD network architecture requires that the Xon2 protocol be carried on UDP packets when delivered over packet-based networks. MPEG2 TS does not, however, perform well over a high-traffic, shared channel such as the public internet where data may travel through multiple network hops comprised of possibly disparate networks.

For channels that suffer lost data from interference Forward Error Correction (FEC) techniques may further ensure adequate performance at the expense of additional overhead. Suitable FEC methods include only those that are transparent (systematic) to receivers not designed to decode the additional protection codes.

#### 5.2.1.2 Interactivity

The delivery of MPEG2 transport stream using UDP/IP over circuit-based or packet-based networks does not afford viewer control to pause, rewind, or skip through the stream, unless the stream is further buffered or saved at the client.

#### 5.2.1.3 Discovery

The discovery of content transported over MPEG2 transport stream requires knowledge of the source; either a frequency to tune if broadcast, or the IP address of the sender if sent via packet-based means.

### 5.2.2 RTP/RTCP

Real Time Protocol (RTP) is a real-time end-to-end transport protocol developed for packet-based networks [7, 8]. While RTP is the carrier of the data, RTCP (Real Time Control Protocol) manages the session by providing feedback on the quality of the data delivered and information about session participants. RTP rides on top of UDP/IP and a RTP session is composed of a RTP port number (UDP port), a RTCP port number (consecutive UDP port) and the participant's IP address.

#### 5.2.2.1 Network Type

RTP is designed for packet-based networks where delivery can not be guaranteed because of packet loss, jitter, and latency. Responsibility for data delivery is at the application level. RTP thus requires smart application endpoints to manage the data for optimal delivery. RTP is a preferred protocol in high-traffic shared networks, but is also applicable to any packet-based network.

Unlike Xon2 that carries all the media components associated with a collect within one package, only one media component is sent per RTP/RTCP connection. For example, video would require one RTP and one RTCP connection, while metadata would require its own separate RTP/RTCP connection. Synchronization of the various streams is re-established at the client through timestamps included in each stream, and a clock reference and timestamp supplied in the respective RTCP channels. Xon2 may be carried over RTP; although this incurs additional bandwidth and is less seldom done.

#### 5.2.2.2 Interactivity

Real Time Streaming Protocol (RTSP) [9] provides for setup, tear-down, and interactive user control of a RTP session. Functions such as pause, rewind, skip, stop, and play are supported by moving through timestamps within the content. This interactivity requires that a streaming server be employed to serve the content. A streaming server is software that formats the content to support interactivity by providing additional information such as hinting tracks. It also helps guard against perturbations incurred in transmission by shuffling the data to non-adjacent packets. Content is organized to support lower resolution and fidelity streams under changing network conditions, like sudden drops in channel bandwidth.

#### 5.2.2.3 Discovery

RTP and RTCP connections are announced (made known) through the guidance of yet other protocols. One commonly used protocol is the Session Announcement Protocol (SAP) which further relies on the Session

Description Protocol (SDP) [10]. SAP periodically issues (announces) IP addresses that clients can listen for and respond to. SDP provides the description of session information for the receivers to connect.

### 5.2.3 HTTP

If real-time delivery is not required the preferred method of delivery is file transfer over HTTP/TCP/IP. HTTP (Hypertext Transfer Protocol) [11] is a network protocol built upon TCP used to connect a web interface to a server. The delivery of data is thus guaranteed by virtue of TCP reliance, and assured transversal through network fire-walls.

#### 5.2.3.1 Network Type

A client requests data by sending both its IP address and port number to port 80 of an addressed server. Port 80 is the default "open" port available for HTTP traffic. This assures that data can get through existing firewalls that might otherwise block reception. The advantage of HTTP is that the data is guaranteed to be delivered, since HTTP uses TCP/IP as the underlying IP protocol. An advantage over RTP, HTTP relies on common web servers to source the motion imagery instead of more costly streaming servers.

#### 5.2.3.2 Interactivity

Two types of HTTP file transfer are common: *progressive download* and *download*. Progressive download provides for file viewing *before* the entire file is downloaded; usually after 10-15% of the file has been received. Media client players must support progressive download to take advantage of this. In *download*, viewing is not available until the entire file is received.

Dependent on the client the user can interact with the content analogous to TiVo. Content is persistent at the client until either a new stream overwrites it, or the player application is terminated. The user has the ability to continually review the content and store the content unlike that for real-time streams.

#### 5.2.3.3 Discovery

Content is accessed through a client web browser by clicking on a contents URL (Uniform Resource Locator). Content is made known through publication of a listing of URL's on a web site.

## 5.3 Metadata

The approved metadata constructs are described in EG0104.6 “Predator UAV Basic Universal Metadata Set and EG0601.1 “UAS Datalink Local Metadata Set.” EG0104.5 specifies a mapping for Cursor on Target (CoT), which can likewise map into EG0601.1 fields. EG0601.1 is preferred as a metadata construct over EG0104.5. Metadata that is synchronized [12] to the video is preferred over asynchronous methods. Security metadata abides by RP 0102.5.

## 5.4 Audio

TBD. The following is under review: *[The approved audio codec is MPEG-4 ISO/IEC 14496-3:2005 High Efficiency-Advanced Audio Coding (HE-AAC) Profile at Level 2 with support for up to 2 channels and 48 kHz sampling. The encoder may use Spectral Band Replication (SBR) to provide a greater frequency range at a given bit-rate, but it is not required. Decoders that cannot process SBR will still be able to decode Advanced Audio Coding (AAC) audio; however, the frequency range will be reduced significantly. Since SBR provides a much higher frequency range at a given bit-rate use of SBR by encoders is recommended.]*

## 5.5 File Format

TBD

## 5.6 Summary

| Motion Imagery for Low Bandwidth Guide | | | |
|---|---|---|---|
| **Media** | **Format** | **Notes** | **References** |
| Motion Imagery Essence | H264/AVC | Maps to Xon2 and RTP | MISP 4.5 |
| Profile | MISM L1.2-L1.0 | | MISP 4.5 |
| Metadata | EG 0601.1, EG 0104.5<br>RP 0102.5 | Mappings to CoT | MISB EG 0601.1<br>MISB EG 0104.5<br>MISB RP 0102.5 |
| Audio | TBD [HE-AAC Profile 2] | Options: 2 channels, SBR | ISO/IEC 14496-3 |
| Transport Protocol | Xon2 | RF, dedicated, or low traffic packet-based networks | MISP 4.5 |
| | RTP/RTCP | Packet-based | |
| File Container | TBD | ISO Base Media Format currently under study | |